Published in final edited form as:

Fungal Genet Biol. 2015 September ; 82: 9–21. doi:10.1016/j.fgb.2015.06.003.

Rapid mapping of insertional mutations to probe cell wall regulation in *Cryptococcus neoformans*

Shannon K. Esher^{a,*}, Joshua A. Granek^{a,b,c,*}, and J. Andrew Alspaugh^{a,d}

^aDepartment of Molecular Genetics & Microbiology, Duke University School of Medicine, Durham, NC 27710, USA

^bDepartment of Bioinformatics and Biostatistics, Duke University School of Medicine, Durham, NC 27710, USA

^cDuke Center for the Genomics of Microbial Systems, Duke University School of Medicine, Durham, NC 27710, USA

^dDepartment of Medicine, Duke University School of Medicine, Durham, NC 27710, USA

Abstract

Random insertional mutagenesis screens are important tools in microbial genetics studies. Investigators in fungal systems have used the plant pathogen *Agrobacterium tumefaciens* to create tagged, random mutations for genetic screens in their fungal species of interest through a unique process of trans-kingdom cellular transconjugation. However, identifying the locations of insertion has traditionally required tedious PCR-based methods, limiting the effective throughput of this system. We have developed an efficient genomic sequencing and analysis method (AIM-Seq) to facilitate identification of randomly generated genomic insertions in microorganisms. AIM-Seq combines batch sampling, whole genome sequencing, and a novel bioinformatics pipeline, AIM-HII, to rapidly identify sites of genomic insertion. We have specifically applied this technique to *Agrobacterium*-mediated transconjugation in the human fungal pathogen *Cryptococcus neoformans*. With this approach, we have screened a library of *C. neoformans* cell wall mutants, selecting twenty-seven mutants of interest for analysis by AIM-Seq. We identified thirty-five putative genomic insertions in known and previously unknown regulators of cell wall processes in this pathogenic fungus. We confirmed the relevance of a subset of these by creating independent mutant strains and analyzing resulting cell wall phenotypes. Through our sequence-based analysis of these mutations, we observed “typical” insertions of the *Agrobacterium* transfer DNA as well as atypical insertion events, including large deletions and chromosomal rearrangements. Initially applied to *C. neoformans*, this mutant analysis tool can be applied to a wide range of experimental systems and methods of mutagenesis, facilitating future microbial genetic screens.

shannon.esh@duke.edu (SK Esher). joshua.granek@duke.edu (JA Granek). Corresponding author: J. Andrew Alspaugh, DUMC 102359, 303 Sands Research Building, Duke University School of Medicine, Durham, NC 27710 USA, andrew.alspaugh@duke.edu (JA Alspaugh).

*These authors contributed equally

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

AIM-Seq; *Cryptococcus neoformans*; *Agrobacterium*-mediated transconjugation; T-DNA; insertional mutagenesis; fungal cell wall; fungal pathogen

1. Introduction

Genetic screens, including those performed by random insertional mutagenesis, have elucidated the functions of many genes in varied species. *Agrobacterium tumefaciens*-mediated transconjugation (AMT) is one method of trans-kingdom DNA delivery that has been increasingly used to generate insertional mutants in a variety of organisms. While *A. tumefaciens* is a plant pathogen in nature, it has been adapted for DNA transformation into a number of species, including several fungi. Fungi that have been studied using AMT include the yeast *Saccharomyces cerevisiae*; the plant pathogens *Fusarium oxysporum*, *Leptosphaeria maculans*, and *Magnaporthe oryzae*; and the human pathogenic fungi *Aspergillus fumigatus*, *Blastomyces dermatitidis*, *Histoplasma capsulatum*, and *C. neoformans* (Betts et al., 2007; Blaise et al., 2007; Bundock et al., 1995; Idnurm et al., 2004; Meng et al., 2007; Mullins et al., 2001; Sugui et al., 2005; Sullivan et al., 2002; Youseff et al., 2009).

The DNA transferred during transconjugation, or T-DNA, is flanked by short border repeats that serve as recognition sequences for processing and transfer by *A. tumefaciens*. When used for AMT, the portion of the T-DNA between the border repeats is replaced with species-specific selectable markers to allow selection of transformed isolates (Michiels et al., 2005). In contrast to other common methods such as transposon mutagenesis, AMT insertions have been reported to occur frequently as single genomic integration events with less bias towards integration at particular loci (de Groot et al., 1998; Michiels et al., 2005).

Cryptococcus neoformans is an opportunistic fungal pathogen that causes over 500,000 deaths per year worldwide primarily in immunocompromised populations (Park et al., 2009). An environmental fungus, *C. neoformans* is inhaled into the lungs, where it establishes a primary infection. In immunocompromised individuals, it can disseminate from this primary site of infection to the central nervous system to cause life-threatening meningitis. *C. neoformans* has emerged as a model organism of fungal pathogenesis, and molecular genetic techniques are well established in this species. Importantly, the genome sequence has recently been published for a clinically derived isolate used widely in many research laboratories (Janbon et al., 2014). AMT has been an important tool for determining the function of a number of genes in *C. neoformans*, including those required for melanin production, sexual development, metal homeostasis, and capsule production (Chun and Madhani, 2010; Feretzaki and Heitman, 2013; Fu et al., 2011; Hu et al., 2013; Idnurm et al., 2004; Lin et al., 2010; Walton et al., 2005).

The process of AMT has been streamlined so that generating and screening large mutant libraries is quite straightforward. However, identifying the sites of AMT insertion continues to be a time- and resource-consuming task. Classical molecular methods, such as inverse PCR, splinkerette PCR, and vectorette PCR often require many rounds of optimization to

identify a limited number of AMT insertion sites (Arnold and Hodgson, 1991; Devon et al., 1995; Leoni et al., 2011; Triglia et al., 1988). The effort required to identify AMT-generated mutations creates a significant bottleneck in an otherwise efficient system. Furthermore, current PCR-based methods regularly fail to identify many induced mutations.

Here we report a new method to rapidly identify sites of insertion generated by techniques such as AMT. AIM-Seq (Agrobacterium-mediated Insertional Mutagenesis sequencing) is a sequencing and analysis method that combines batch sampling, whole genome sequencing, and a new bioinformatics pipeline, AIM-HII (Agrobacterium-mediated Insertional Mutagenesis High-throughput Insert Identification), to identify the sites of insertion in mutants generated by AMT. This method has several key advantages over established PCR-based identification methods, as well as other insertion sequencing methods such as Tn-Seq or INSeq (Barquist et al., 2013; Goodman et al., 2009; van Opijnen et al., 2009). Previously generated AMT libraries can be used with no additional steps or optimization required. This is in contrast to transposon-insertion sequencing methods that require additional steps to enrich for insert-flanking fragments prior to sequencing (Barquist et al., 2013). Although higher throughput than traditional PCR-based insertion identification methods, these previously described deep sequencing methods still pose technical challenges especially biases introduced by initial PCR-enrichment steps. While applicable in theory to eukaryotes, many of these established techniques were designed with smaller, prokaryotic genomes in mind. Additionally, AIM-Seq is capable of identifying non-canonical insertion events that cannot be assessed by standard PCR-based and insertion sequencing methods. Furthermore, when considering the overall cost, including labor, of performing PCR-based identification methods for each mutant selected, the cost of the DNA extraction and whole genome sequencing for AIM-Seq is less expensive on a per mutant basis.

The *C. neoformans* cell wall plays a key role in immune recognition and avoidance. Given this important cellular function, we sought to characterize mutants with changes in the cell wall that may affect its interaction with the host. We screened a library of random AMT mutants for altered cell wall phenotypes, applying the AIM-Seq method to map and confirm the mutations. In this screen we identified “typical” T-DNA insertions in genes known to be involved in cell wall homeostasis, as well as in genes not previously associated with the fungal cell wall. We also identified atypical AMT-induced events that would be missed by traditional mutation identification methods. We believe that these types of events are responsible for previously reported mutations of interest that could not easily be identified by standard means (Feretzaki and Heitman, 2013; Fu et al., 2011; Hu et al., 2013).

While the proof of concept experiments discussed here were conducted in *C. neoformans* using AMT generated mutants, this method is applicable more generally: it can be used with any insertional mutagenesis method (e.g. transposon-mediated mutagenesis) and in any mutable target organism. The only requirements are that the insert sequence and target genome sequences are known (or can be sequenced). Once mutants are generated and screened, organism-specific adjustments to this analysis tool can be easily made by choosing a sequencing depth appropriate to the genome size and the number of mutations expected.

2. Materials and Methods

2.1 Strains, media, and growth conditions

Cryptococcus neoformans strains used in this study are listed in Table 1. *A. tumefaciens* strain EHA105/NAT was used for AMT (Walton et al., 2005). Unless otherwise noted, all strains were created in the *C. neoformans* H99 *MAT α* background (Janbon et al., 2014). Strains were cultured on YPD (yeast extract 1%, peptone 2%, dextrose 2%) agar plates or in YPD liquid media with shaking at 150 rpm at 30°C, unless otherwise stated (Sherman, 1991). To induce capsule for India ink visualization, cells were incubated in CO₂-independent tissue culture media (Gibco) for 3 days with shaking at 150 rpm at 37°C. Melanin was assessed at 30°C and 37°C on Niger seed (*Guizotia abyssinica*) medium prepared as described previously (Kwon-Chung and Bennett, 1992). Congo red (0.5%) was added to YPD medium prior to autoclaving. Caffeine (1 mg/mL) and calcofluor white (1 mg/mL) were filter sterilized and added to YPD after autoclaving. Quinolinic acid (Sigma Aldrich) was filter sterilized and spread onto yeast nitrogen base (YNB) plates at the indicated concentrations.

For cell staining and imaging, seed cultures were grown overnight in YPD medium with shaking at 150 rpm at 30°C. Cells were then diluted into YPD or CO₂-independent tissue culture medium and incubated with shaking at 150 rpm at 30°C or 37°C for 18 hours.

2.2 Molecular Biology

All primers used in this study are listed in Table 2. Independent confirmatory mutants were generated by amplifying the insert T-DNA cassette from the original AMT mutant, followed by genomic integration by biolistic transformation into the H99 *MAT α* background (Toffaletti et al., 1993). All new deletion strains were confirmed by Southern blot analysis (data not shown), using the nourseothricin (NAT) resistance gene as a probe (McDade and Cox, 2001).

2.3 *Agrobacterium tumefaciens*-mediated transformation (AMT)

AMT was carried out as previously described using *A. tumefaciens* strain EHA105 expressing the PZP-NATcc plasmid (Idnurm et al., 2004; Walton et al., 2005). In brief, *A. tumefaciens* was incubated overnight with shaking at 30°C in Luria-Bertani (LB) media containing kanamycin. *A. tumefaciens* cells were washed and resuspended at an OD_{660nm} of 0.15 in liquid induction medium (IM) with 100 μ M acetosyringone (AS), shaking at room temperature for 6 hours, or until the final concentration reached an OD_{660nm} of 0.6 (Bundock et al., 1995). An overnight culture of *C. neoformans* H99 was harvested and diluted in IM to a final concentration of 10⁶ cells/mL. Equal volumes (200 μ L) of *A. tumefaciens* and *C. neoformans* cells were then co-incubated on IM agar at room temperature for 2 days, after which cells were scraped onto selection media (YPD + 100 μ g/mL nourseothricin + 100 μ g/mL cefotaxime). Colonies were inoculated into liquid YPD media in 96-well plates and incubated at 30°C for 2-3 days before freezing down and/or screening. A wild type H99 control was inoculated into wells A1 and A7 of each 96-well plate.

2.4 Screening of AMT mutants

1500 mutants were patched onto YPD agar media containing 150 mM HEPES adjusted to pH 8, 1.5 M NaCl, or 0.06% sodium dodecyl sulfate (SDS). Mutants with one or more growth phenotypes on these conditions were selected and passaged 1x on YPD and inoculated into liquid YPD media in new 96-well plates, followed by re-patching for confirmation.

2.5 Preparation of genomic DNA and whole genome sequencing

Forty AMT mutants were selected for further genotypic analysis. We made our initial selection of 40 mutants based on the size of the *C. neoformans* genome and the expected sequencing output of the MiSeq sequencing platform. The number of mutants pooled can be adjusted to fit the genome size of the organism and sequencing platform. We estimated that with these conditions we would obtain an average genome coverage of approximately 10x for each of the 40 strains. Genomic DNA was extracted from each mutant using the MasterPure Yeast DNA Purification Kit (Epicentre). The number of insertions in each mutant was assessed by Southern blot using a probe for the heterologous NAT resistance gene (Table 2). The genomic DNA of 27 AMT mutants with single insertions was pooled in equimolar amounts to a final concentration of 40 ng/μL. This pooled DNA was submitted to the Duke University Genome Sequencing and Analysis Core for library preparation and sequencing (Illumina MiSeq, 250 bp × 250 bp, paired end reads). The resulting sequence data consisted of 1.75×10^7 read pairs, yielding an average of 8.5× genomic coverage per strain.

2.6 Identification of inserts by AIM-HII

The AIM-Seq analysis pipeline can be accessed at <https://github.com/granek/aimhii>. AIMHII, the analysis software, is a Python pipeline that uses several pre-existing packages as well as implementing methods developed specifically for this analysis (Supplementary figure 1). The required input to AIM-HII includes a reference genome sequence, insert DNA sequence, adapter sequences, and sequence read data in FASTQ format (either single-end or paired-end). The appropriate adapter sequence was supplied by Illumina (TruSeq Adapter Index 1:

GATCGGAAGAGCACACGTCTGAACTCCAGTCACATCACGATCTCGTATGCCGTC
TTCTGCTTG, Oligonucleotide sequences © 2007-2013 Illumina, Inc. All rights reserved).

There are a number of adjustable parameters that can be provided when running AIM-HII, including the maximum allowable insertion-site deletion, the minimum number of reads required for a cluster, and the number of CPUs that should be utilized for parallel operation.

AIM-HII removes sequencing adapters and trims low quality bases using fastq-mcf (Aronesty, 2013). If paired-end data is supplied, read pairs with overlapping 3' ends are merged into a single read using fastq-join (Aronesty, 2013). AIM-HII concatenates the genome and insert sequence supplied by the user, then runs BWA-MEM (Li, 2013) for mapping the processed reads to this concatenated reference sequence. It converts the SAM file output by BWA-MEM to a sorted BAM file using samtools (Li et al., 2009). AIM-HII then extracts all reads spanning a genome-insert junction. These reads are processed to determine where in the genome the putative insertion occurred, and then assembled into

clusters by their genomic location. AIM-HII pairs clusters that are within the specified gap limit and appear to flank opposite ends of the insert into “cluster pairs” (Figure 2B). We have found that a subset of read clusters have no identifiable partner; we refer to these unpaired clusters as “singleton clusters” (Figure 2D). AIM-HII uses a conservative default maximum gap (i.e., insertion-site deletion size) of 5 kb but this value can be adjusted at run time, to accommodate insertional systems with different insertion site deletion profiles. Because neither individual reads, nor read pairs can span a full insertion, insert size and coordinates are extrapolated based on the insert portions of the junction spanning reads. The outputs from AIM-HII include a table with detailed information about each cluster pair or singleton cluster, and a plot of the constituent reads. AIM-HII depends on a number of non-standard Python libraries: Pysam (Heger et al., 2014), HTSeq (Anders et al., 2014), NumPy (van der Walt et al., 2011), and Biopython (Cock et al., 2009).

2.7 Detailed phenotypic analysis of mutant strains

To further characterize newly identified genes, we performed detailed phenotypic analyses of the original (AMT mutant) and independent mutant strains. We directed our phenotypic analysis to include cellular features associated with pathogenesis, including growth at 37°C, capsule and melanin production, as well as additional cell wall-associated phenotypes. We examined the susceptibility of our mutants to cell wall disrupting agents including Congo red, caffeine, and calcofluor white, and performed cell wall staining for exposed chitin and chitosan (see below).

2.8 Microscopy

All differential interference contrast (DIC) and fluorescent microscopy images were captured using a Zeiss Axio Imager A1 fluorescent microscope equipped with an AxioCam MRM digital camera.

2.9 Cell wall staining

Similar to other cell wall components, chitin has been shown to be recognized by the host immune system as a pathogen associated molecular pattern (Bueter et al., 2013; Goldman and Vicencio, 2012; Vega and Kalkum, 2012; Wagener et al., 2014). The *C. neoformans* cell wall in particular contains significantly more chitin and chitosan than other pathogenic fungi (Banks et al., 2005). Our lab has previously shown an even greater increase in these molecules in the highly immunogenic *rim101* mutant strain cell wall, specifically when these cells are grown in tissue culture conditions (O’Meara et al., 2014, 2013). For these reasons, we chose specifically to examine the exposure of chitooligomers in our cell wall mutants.

To visualize chitin, cells were pelleted, washed twice with PBS, and stained with 100 µg/mL wheat germ agglutinin (WGA) conjugated to Alexa Fluor 488 (Molecular Probes) for 30 minutes at room temperature, followed by 25 µg/mL CFW (source) for 10 minutes at room temperature. Cells were washed twice with PBS prior to imaging. WGA staining was visualized using a GFP filter and CFW staining was visualized using a DAPI filter.

To visualize chitosan, cells were stained with eosin Y as described previously (Baker et al., 2007). Cells were pelleted and washed twice with McIlvaine's buffer (0.2 M Na₂HPO₄ and 0.1 M citric acid, pH 6.0), followed by staining with 300 µg/mL eosin Y (EY, Sigma) for 5 minutes at room temperature. Cells were then washed twice more with McIlvaine's buffer prior to imaging. Eosin Y staining was visualized using a GFP filter.

2.10 FM4-64 staining

Cells were pelleted, washed twice with PBS, and stained with a final concentration of 2 µg/mL FM4-64 (Molecular Probes) for 10 minutes on ice. Cells were then washed with PBS, resuspended in fresh PBS, and incubated at room temperature for 30 minutes prior to imaging. FM4-64 was visualized using a Texas Red filter.

2.11 Software and Data Availability

AIM-HII is open source software and is available for download. The AIM-HII repository at <https://github.com/granek/aimhii> contains instructions for installing and running AIM-HII, and the AIM-HII source code. An AIM-HII Python package is at <https://pypi.python.org/pypi/aimhii>, and an AIM-HII Docker image is at <https://registry.hub.docker.com/u/granek/aimhii/>. This Docker image allows AIM-HII to be run on a number of computer types including Mac OS X, Windows, and Linux. The raw sequences from our pooled mutants have been deposited at the Sequence Read Archive (Leinonen et al., 2011); the SRA Study accession number is SRP057031. In keeping with best practices of reproducible research (Peng, 2009), the software repository also includes a script, support files, and instructions for automatically replicating all analyses presented here using the data available from the Sequence Read Archive.

3. Results

3.1 Insertional mutagenesis screen for *C. neoformans* cell wall mutants

We screened a library of 1500 *C. neoformans* insertional mutants for cell wall phenotypes known to be associated with adaptation to the host environment (O'Meara et al., 2013, 2010). These phenotypes included sensitivity to pH 8, sensitivity to high concentrations of salt (1.5 NaCl), and resistance to the detergent sodium dodecyl sulfate (0.06% SDS). We selected 40 mutants that exhibited one or more of these phenotypes.

Previous investigations suggest that AMT in *C. neoformans* primarily results in single, randomly generated genomic insertion events with rare instances of multiple insertions (Fu et al., 2011; Hu et al., 2013; Idnurm et al., 2009). To simplify our initial assessment process, we limited our analysis to strains with a single insertion. Our Southern blot analysis identified 27 of 40 mutant strains with a single likely insertion of the nourseothricin resistance marker (NAT) that is part of the T-DNA (data not shown).

3.2 Mutant insert identification by *Agrobacterium* Insertional Mutagenesis Sequencing (AIM-Seq)

In order to facilitate the identification of random insertion events generated by AMT, we developed AIM-Seq, a high-throughput genomic sequencing and analysis method (Figure

1). Genomic DNA from each of the 27 single insertion mutants was extracted and pooled in equimolar amounts. This pooled DNA sample was analyzed by whole genome sequencing. We sequenced the pooled genomic DNA in a single MiSeq lane with 250 bp $\times$ 250 bp paired-end reads (Illumina, Inc.). With 250 bp reads, it is not possible for a single read to fully span the expected insert, which is over 2 kb long. Instead, our strategy to map insertion locations is to identify the sequences flanking each insertion by finding reads that span genome-insertion junctions. Long reads are particularly important in this application because the longer the read, the higher likelihood that this particular sequence will span a genome-insertion junction and contain enough sequence on each side of the junction that both parts can be uniquely mapped.

This sequencing data was next analyzed by AIM-HII (*Agrobacterium* Insertional Mutagenesis High-throughput Insert Identification), the software we developed to identify sites of insertion from our whole genome sequencing data (Supplementary Figure 1). AIM-HII takes as input the raw whole-genome sequence data from the mutant pools; maps and extracts reads that contain both genome and insert sequence; clusters them by their genomic location; and identifies the pair of read clusters that flank a putative insertion. The AIM-HII output is formatted as a table with detailed information about identified insertions as well as plots of the constituent reads (Supplementary Table 1). Additionally, the AIM-HII dataset can be examined directly using genomic visualization tools such as the Integrative Genomics Viewer (IGV) (Robinson et al., 2011; Thorvaldsdóttir et al., 2013).

3.3 AIM-HII identified typical and atypical insertion events

In our mutant strain collection, AIM-HII identified 35 putative insertion sites with greater than 5 total aligning sequences (Table 3, Supplementary Table 1, Figure 2). Sixteen of these sites displayed typical insertion characteristics with clustered reads on both sides of the putative insertion site (Figure 2B). Of these typical insertions, 10 displayed short deletions of 2-30 bp, similar to what has been observed previously using this mutagenesis technique in *C. neoformans* (Idnurm et al., 2004; Shimizu et al., 2014). Four sites displayed no deletion, or a putative small DNA duplication at the site of T-DNA insertion. The remaining two events involved larger insert-induced deletions, including one deletion of 1.2 kb (Figure 2C). Large deletions have been described in other fungi, however they are thought to occur uncommonly (Kemski et al., 2013; Li et al., 2007; Meng et al., 2007).

At the remaining 19 putative insertion sites, we observed atypical phenomena we have named “singleton clusters”, in which only one side of a putative insertion site has true clustering reads (Figure 2D). We hypothesize that these singleton clusters represent insertion-induced large chromosomal rearrangements. These types of events have been documented in AMT mutants in *C. neoformans* and other organisms (Gheysen et al., 1991; Idnurm et al., 2009; Kemski et al., 2013; Li et al., 2007; Michielse et al., 2005).

3.4 Insertion-induced alterations of the T-DNA sequence

The AIM-Seq method also gives us detailed insight into the variable structure of the T-DNA. Five of the 35 putative sites have truncated T-DNA inserts, as indicated by insert lengths significantly shorter than 2.3 kb, the full T-DNA sequence length (Supplementary

Table 1). In all of the truncated inserts in our experiments, the size of the insert sequence suggests that part or all of the selectable marker is missing. Since all the AMT strains were initially selected for their NAT resistance, it is likely that these truncated inserts are additional mutation events occurring in strains with “typical” insertions of the selectable marker. It would not be immediately evident whether the observed phenotypes of these strains are due to their “typical” NAT containing insert or their truncated insert; however, with both sites of insertion identified, independent re-creation of these mutations would distinguish which insertion was causative. In addition, these truncated insertions will be missed by PCR-based methods, including transposon-insertion sequencing methods (Barquist et al., 2013), which depend on specific sequence being present near the end of the inserted sequence. Therefore, our whole genome sequencing-based methods of insert localization identified nontraditional mutations in our strains, which could not have been identified by established methods.

Truncations, as well as extensions, of the T-DNA insert have been reported previously in AMT experiments from diverse fungi (Bundock et al., 1995; Idnurm et al., 2009; Kemski et al., 2013; Meng et al., 2007; Michielse et al., 2005). Our identification of vector backbone beyond the T-DNA insert sequence indicates that, at least in one case, additional and unexpected vector sequence was introduced into the fungal strain (Figure 3). While incorporation of the vector backbone is theoretically identifiable by PCR-based methods, it would complicate the interpretation of the results, as the PCR product would be much larger than expected. These data confirm that AMT-generated mutants in fungal species do not always contain simple, clean insertions of the selectable marker, and represent another group of mutants that might be difficult to characterize by existing methods.

3.5 Confirmation of AIM-HII putative insertion sites by PCR

We confirmed the output of AIM-HII by PCR, by designing primers on both sides of putative insertion sites to amplify across the insert. In this manner, we were able to distinguish between wild type and mutated loci by differences in PCR fragment size. By PCR screening all of the mutants, we have matched 9 insertions to individual strains, indicating that AIM-HII positively identified these true insertion sites (Table 4). We attempted to amplify across 3 separate singleton events, however we were unable to generate a corresponding PCR amplicon. Considering these events may be due to chromosomal rearrangements, it was expected that they could not be identified in this manner. While this site-to-strain matching was used to validate the AIM-Seq method, it is not a necessary step in *C. neoformans*. In this organism where both classical and reverse genetics are possible, identification of the insertion site will allow the simple re-creation of the relevant mutant strains, either with a complete gene deletion, or with an identical mutation as in the original insertional mutant. Importantly, a new mutant strain will also allow independent confirmation of the mutant and its phenotypes.

3.6 Phenotypic validation of AIM-Seq method

In order to further validate the insert identification method, we tested whether the phenotypes of the original AMT mutants were identical in independently generated mutant strains. Table 5 details the mutations that have been phenotypically validated with

independent mutant strains. Three of the mutations identified by AIM-Seq occurred in previously described genes, for which pre-existing mutant strains with targeted deletions of these genes were available.

One of our new mutants, WHM24, contained an insertion in exon 6 of the gene encoding adenylyl cyclase-associated protein 1 (*ACA1*, CNAG_05218). This mutation resulted in a strain exhibiting a growth defect at pH 8 and in the presence of high salt (Figure 4). A previously generated *aca1* mutant strain exhibited similar pH and salt sensitivities (Bahn et al., 2004). Previous work has demonstrated that Aca1 interacts with the adenylyl cyclase Cac1 to regulate capsule and melanin (Bahn et al., 2004). Phenotypic analysis of the *aca1* mutant compared to other cAMP/PKA pathway mutants revealed a role for this protein in cell wall integrity (Maeng et al., 2010), confirming the ability of our screen to identify mutants with important cell wall phenotypes.

Another of our new mutants, WHM21, contained an insertion in the 5' untranslated region of the *SKN7* gene encoding a response regulator and transcription factor (CNAG_03409). The original WHM21 strain exhibited salt sensitivity and resistance to SDS (Table 5). An independently derived *skn7* mutant has been previously described to be salt sensitive, and it was recently shown to be at least 10-fold more resistant to SDS than WT (Bahn et al., 2006; Coenjaerts et al., 2006; Jung et al., 2015; Wormley et al., 2005). The *SKN7* gene was originally identified in *S. cerevisiae* through a screen to identify proteins that suppress mutations in *kre9* mutants, which have decreased cell wall β -1,6-glucan (Brown et al., 1993). Interestingly, the AMT-induced mutation in WHM21 also affected the potential regulatory/protein-coding region of a neighboring gene (CNAG_03410). However, the phenotypic similarity between the independently derived *skn7* mutant strain and WHM21 suggests that the WHM21 mutant cell wall phenotypes are primarily due to altered *SKN7* expression.

Our mutant phenotype analyses indicate that at least one of our identified insertion sites was not the cause of the cell wall phenotype observed in the corresponding mutant strain. We identified a simple insertion in the first exon of an arrestin-like gene, *ALII* (CNAG_02857) in mutant strain WHM13. An independently created *ali* mutant (Ost et al., 2015) did not display any of the cell wall-associated phenotypes that we tested. Since more mutations were identified than the total number of mutant strains tested, it would be expected that not all identified T-DNA insertion events were necessarily causative for mutant phenotypes of interest. Furthermore, previous studies have shown through linkage analysis that AMT mutant phenotypes are not always associated with the T-DNA insertion event (Feretzi and Heitman, 2013; Idnurm et al., 2004; Michielse et al., 2005; Walton et al., 2005), emphasizing the need for subsequent biological validation of identified mutations.

3.7 Identification of new genes with cell wall related functions

We also generated several independent mutant strains to confirm a role in cell wall integrity for genes that have not been previously characterized. In each case we constructed an independent mutant by transforming the WT strain with a disruption construct generated by PCR-amplifying insert-containing fragments from the mutant pool. In this way, we confirmed that these mutations in three previously uncharacterized genes were responsible

for the cell wall phenotypes observed in the associated mutant strains (Table 5). These genes include a putative mannosyltransferase (CNAG_05142, WHM4), a putative kynurenine-3-monooxygenase (CNAG_02793, WHM11), and a conserved hypothetical protein (CNAG_02676, WHM27).

The CNAG_05142 gene locus encodes a predicted α -1,3-mannosyltransferase. The *Saccharomyces cerevisiae* homolog of this gene, *ALG3*, encodes an α -1,3-mannosyltransferase that initiates one of the branches of *N*-glycan from a dolichol pyrophosphate backbone in the beginning stages of protein *N*-glycosylation (Aebi et al., 1996; Sharma et al., 2001). Both the original and independent mutant strains grew well at 30°C and 37°C and displayed no defects in melanin production; however, a slight decrease in capsule production was observed in both mutants (Figure 5). These mutant strains displayed growth inhibition on 1.5 M NaCl, but did not display altered phenotypes on pH 8 or 0.06% SDS (Table 5). Additionally, the strain does not have any growth defects on caffeine, calcofluor white (CFW), or Congo red, nor does it display differences from WT in cell wall staining, suggesting no alterations in exposed chitin or chitosan content (Figure 5, 6). Together these results suggest that this gene, although identified in our primary screen for regulating salt sensitive growth, may not be involved in major cell wall-associated processes.

CNAG_02793 encodes a predicted kynurenine-3-monooxygenase (KMO) enzyme. Our original and independent mutant strains demonstrated growth defects at 37°C, as well as a defect in capsule (Figure 5). These mutants also displayed modest sensitivities to CFW and Congo red (Figure 5). Staining with CFW and WGA revealed increased fluorescence when the cells were grown in tissue culture conditions, suggesting an increase in chitin exposure under these conditions (Figure 6, Supplementary Figure 2). This predicted KMO enzyme is a member of the kynurenine pathway that converts tryptophan to quinolinic acid (QA), which in yeast is utilized to produce NAD⁺. The kynurenine-3-monooxygenase enzyme, encoded by the gene *BNA4* in *S. cerevisiae*, catalyzes an intermediate step in this process, converting kynurenine to 3-hydroxykynurenine (3-HK) (Ohashi et al., 2013; Panozzo et al., 2002). We observed that these mutants were unable to grow on YNB (yeast nitrogen base) media; however we were able to suppress this growth defect by adding exogenous QA, demonstrating that this protein is likely functioning as the KMO enzyme (Bna4) in *C. neoformans* (Figure 7). The kynurenine pathway has not previously been linked to changes in the fungal cell wall.

Finally, CNAG_02676 encodes a conserved hypothetical protein with homology to a vacuolar SNARE protein in *S. cerevisiae*, *VAM7*, which has been implicated in vacuolar transport and morphogenesis (Sato et al., 1998; Ungermann and Wickner, 1998; Wada and Anraku, 1992; Wada et al., 1992). A modest decrease in capsule production was observed in the original and independent mutants, as well as attenuated melanin production at both 30°C and 37°C, despite wild type growth at all temperatures (Figure 5). CFW and WGA staining of the mutants were increased under both YPD and tissue culture conditions, while eosin Y staining was increased only under tissue culture conditions, suggesting alterations in both chitin and chitosan in the cell wall in host-mimicking growth conditions (Figure 6, Supplementary Figure 2). To assess the role of this gene in *C. neoformans* vacuolar

morphogenesis, we stained wild type and mutant cells with the endocytosis marker, FM4-64. Similar to what has been observed in *S. cerevisiae*, we observed FM4-64 staining of numerous small punctate structures in the mutants, as compared to vacuolar staining in the wild type (Wada and Anraku, 1992). This staining is consistent with a role for CnVam7 (CNAG_02676) in assisting endocytic transport from the membrane to the vacuole. Previous work characterizing this protein in *S. cerevisiae*, as well as homologous proteins in other fungi, have indicated additional roles for Vam7 in fungal cell wall stress and organization (Dou et al., 2011; Lesage et al., 2005; Wedlich-Söldner et al., 2000; Zhang et al., 2015).

4. Discussion

The fungal cell wall is a critical regulator of the host-pathogen interface, representing a dynamic structure that changes in response to different environments. *C. neoformans* dramatically alters its cell wall in response to the host, an important part of this pathogen's ability to adapt to this potentially hostile environment. Here we screened a library of insertional mutants for phenotypes associated with host environment-induced cell wall changes. In our initial screen, we identified two strains with mutations in genes known to regulate cell wall processes in *C. neoformans*. The *ACA1* gene encodes a protein that coordinately signals with the cAMP/PKA pathway. Similar transcriptome patterns were observed between *aca1* and *cac1* (adenylyl cyclase) mutants, supporting an interaction between these two proteins (Maeng et al., 2010). A small group of genes was found to be differentially and specifically regulated in the *aca1* mutant, suggesting a cAMP/PKA-independent downstream signaling output for Aca1 (Maeng et al., 2010). Additionally, this mutant strain has been previously linked to cell wall phenotypes, confirming the screening methodology. Similarly, we identified an insertion in the *SKN7* gene by our screening strategy. While the *skn7* mutant is not sensitive to traditional cell wall stressors such as Congo red or calcofluor white, transcriptional profiling of the *skn7* mutant in *C. neoformans* revealed that approximately 6% of up-regulated genes in this strain compared to the WT were related to cell wall/membrane/envelope biogenesis (Jung et al., 2015; Ko et al., 2009).

We also identified two new genes in *C. neoformans* that have not been previously associated with cell wall processes. *BNA4* (CNAG_02793) encodes a predicted kynurenine-3-monooxygenase involved in tryptophan degradation. The corresponding mutant displayed increased chito oligomer exposure on the cell surface, a new phenotype not previously attributed to this pathway. Similarly, the *VAM7* gene (CNAG_02676) encodes a conserved protein that is involved in endocytosis and vacuolar morphogenesis. Vesicle transport to and from the cell surface plays a major role in cell wall maintenance and the delivery of proteins to the cell surface (Rodrigues et al., 2008a, 2008b; Wolf et al., 2014).

To facilitate our screen, we developed a high-throughput genomic sequencing and analysis method for rapid identification of random insertion events generated by AMT. AIM-Seq is designed to complement the use of lower throughput, PCR-based methods in screening through a large collection of randomly generated mutant strains. It is also a distinct alternative to other deep-sequencing, insert-identification techniques previously used in prokaryotic systems that rely on pre-amplification of transposon-genome junctions prior to

sequencing (Gawronski et al., 2009; Goodman et al., 2009; Langridge et al., 2009; van Opijnen et al., 2009).

In this initial group of experiments, we have established the basic AIM-Seq technique; however, we took several steps in our analysis of this initial mutant set that would not be necessary in routine use. Based on previous studies in which a small portion of AMT mutants contained multiple T-DNA insertion events, we prescreened for mutants that contained only a single insertion. We also took the additional step of site-to-strain matching in order to demonstrate the efficacy of our system. In *C. neoformans* and other organisms where classical and reverse genetics manipulations are possible, and in which AIM-Seq insertion sites will be confirmed through the generation of independent mutants, these more time-intensive steps will generally not be necessary.

In addition to rapidly and inexpensively identifying sites of insertion in our pilot mutant pool, AIM-Seq also allowed us to observe a number of atypical AMT-induced insertion events. Many of these have been suggested or described by previous investigators, both in plants and fungi (Gheysen et al., 1991; Idnurm et al., 2009; Kemski et al., 2013; Li et al., 2007; Mayerhofer et al., 1991; Michielse et al., 2005). Our results suggest that these atypical events might be responsible for mutant phenotypes reported in prior studies in which traditional methods failed to identify the causative mutation. Truncations of the T-DNA are particularly problematic, as we observed that these events can result in the loss of primer binding sites used in many of the PCR-based methods, including transposon-insertion sequencing, thus rendering these sites undetectable by such methods. Therefore, it would be difficult to directly compare transposon- insertion sequencing processes that require intact genome-insert borders with AIM-Seq. Tolerance of truncated inserts is an advantage of AIM-Seq, particularly in mapping insertions in mutants identified through labor-intensive screens where every mutant is precious.

Additionally, while we selected for strains that we believed to carry single insertions, we have putatively identified an AMT strain with two independent insertions; the second insertion was undetectable by Southern blot analysis due to a truncation that eliminated the portion of T- DNA that was used as a probe (data not shown). We hypothesize that the singleton clusters we observed are the result of large chromosomal events such as rearrangements. In fact, detailed analysis of AMT insertion events in *H. capsulatum* revealed that up 31% of characterized T- DNA integrations were associated with genomic rearrangements (Kemski et al., 2013). In *C. neoformans*, Idnurm *et al.* similarly detected both large-scale genomic rearrangements and intrachromosomal rearrangements associated with T-DNA integrations (Idnurm et al., 2009). In our case, these rearrangements likely produce strains containing two mapped singleton events, each corresponding to one side of the mutation. Since each singleton event is identified independently by AIM-Seq, it is difficult to predict which are paired within a given strain. Even if located, the multifactorial effects of such rearrangements (e.g. gene disruption, promoter exchange, and position effects) make it difficult to determine the gene(s) responsible for the phenotype of interest, making these less desirable mutations to study in AMT experiments. The prevalence of these atypical AMT insertion events underscores the need for independent confirmatory studies of AMT generated mutant strains.

Despite the increasing availability of whole gene knockout collections in model organisms, random insertional mutagenesis remains a valuable research tool. Because random insertional mutagenesis libraries can be generated easily and inexpensively, this method enables genetic study of a large diversity of organisms, including wild type isolates and species that do not have established genetic techniques. Insertional mutagenesis is also capable of generating diverse mutations beyond what is available in targeted gene knockout collections, including partial loss-of-function and gene regulatory variants. As such AIM-Seq is an important advancement in leveraging high-throughput sequencing technology in order to make the analysis of insertional mutagenesis more accessible, faster, and cost effective.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by National Institutes of Health grant AI074677 (JAA). We would like to thank the Woods Hole Marine Biological Laboratory Molecular Mycology Class of 2012 for assisting in the creation of the *C. neoformans* insertional mutagenesis library used in this study. We also thank Sandra Breeding for strain curation. We would like to thank the Duke University Genome Sequencing and Analysis Core Resource for Illumina library preparation and sequencing.

Abbreviations

AMT	<i>Agrobacterium</i> -mediated transconjugation
AIM-HII	<i>Agrobacterium</i> Insertional Mutagenesis High-throughput Insert Identification
AIM-Seq	<i>Agrobacterium</i> Insertional Mutagenesis Sequencing
AS	acetosyringone
CDS	coding DNA sequence
CFW	calcofluor white
DIC	differential interference contrast
EY	eosin Y
IGV	Integrative Genomics Viewer
IM	induction medium
NAT	nourseothricin
SDS	sodium dodecyl sulfate
T-DNA	transfer DNA
WGA	wheat germ agglutinin

References

Aebi M, Gassenhuber J, Domdey H, te Heesen S. Cloning and characterization of the *ALG3* gene of *Saccharomyces cerevisiae*. *Glycobiology*. 1996; 6:439–44. [PubMed: 8842708]

- Anders S, Pyl PT, Huber W. HTSeq – A Python framework to work with high-throughput sequencing data. *bioRxiv*. 2014:1–5.
- Arnold C, Hodgson IJ. Vectorette PCR: a novel approach to genomic walking. *Genome Res*. 1991; 1:39–42. doi:10.1101/gr.1.1.39.
- Aronesty E. Comparison of Sequencing Utility Programs. *Open Bioinforma. J*. 2013; 7:1–8. doi: 10.2174/1875036201307010001.
- Bahn Y, Hicks JK, Giles SS, Gary M, Heitman J, Cox GM. Adenylyl Cyclase-Associated Protein Aca1 Regulates Virulence and Differentiation of *Cryptococcus neoformans* via the Cyclic AMP-Protein Kinase A Cascade. *Eukaryot. Cell*. 2004; 3:1476–1491. doi:10.1128/EC.3.6.1476. [PubMed: 15590822]
- Bahn Y, Kojima K, Cox GM, Heitman J. A unique fungal two-component system regulates stress responses, drug sensitivity, sexual development, and virulence of *Cryptococcus neoformans*. *Mol. Biol. Cell*. 2006; 17:3122–3135. doi:10.1091/mbc.E06. [PubMed: 16672377]
- Baker LG, Specht C. a, Donlin MJ, Lodge JK. Chitosan, the Deacetylated Form of Chitin, Is Necessary for Cell Wall Integrity in *Cryptococcus neoformans*. *Eukaryot. Cell*. 2007; 6:855–67. doi:10.1128/EC.00399-06. [PubMed: 17400891]
- Banks I, Specht C, Donlin M, Gerik K, Levitz S, Lodge J. A Chitin Synthase and Its Regulator Protein Are Critical for Chitosan Production and Growth of the Fungal Pathogen *Cryptococcus neoformans*. *Eukaryot. Cell*. 2005; 4:1902–1912. doi:10.1128/EC.4.11.1902. [PubMed: 16278457]
- Barquist L, Boinett CJ, Cain AK. Approaches to querying bacterial genomes with transposon-insertion sequencing. *RNA Biol*. 2013; 10:1161–9. doi:10.4161/rna.24765. [PubMed: 23635712]
- Betts MF, Tucker SL, Galadima N, Meng Y, Patel G, Li L, Donofrio N, Floyd A, Nolin S, Brown D, Mandel MA, Mitchell TK, Xu J-R, Dean R. a, Farman ML, Orbach MJ. Development of a high throughput transformation system for insertional mutagenesis in *Magnaporthe oryzae*. *Fungal Genet. Biol*. 2007; 44:1035–49. doi:10.1016/j.fgb.2007.05.001. [PubMed: 17600737]
- Blaise F, Rémy E, Meyer M, Zhou L, Narcy J-P, Roux J, Balesdent M-H, Rouxel T. A critical assessment of *Agrobacterium tumefaciens*-mediated transformation as a tool for pathogenicity gene discovery in the phytopathogenic fungus *Leptosphaeria maculans*. *Fungal Genet. Biol*. 2007; 44:123–38. doi:10.1016/j.fgb.2006.07.006. [PubMed: 16979359]
- Brown JL, North S, Bussey H. *SKN7*, a yeast multicopy suppressor of a mutation affecting cell wall beta-glucan assembly, encodes a product with domains homologous to prokaryotic two-component regulators and to heat shock transcription factors. *J. Bacteriol*. 1993; 175:6908–15. [PubMed: 8226633]
- Bueter CL, Specht C. a, Levitz SM. Innate Sensing of Chitin and Chitosan. *PLoS Pathog*. 2013; 9:e1003080. doi:10.1371/journal.ppat.1003080. [PubMed: 23326227]
- Bundock P, den Dulk-Ras A, Beijersbergen A, Hooykaas PJ. Trans-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. *EMBO J*. 1995; 14:3206–14. [PubMed: 7621833]
- Chun CD, Madhani HD. Ctr2 Links Copper Homeostasis to Polysaccharide Capsule Formation and Phagocytosis Inhibition in the Human Fungal Pathogen *Cryptococcus neoformans*. *PLoS One*. 2010; 5. doi:10.1371/journal.pone.0012503.
- Cock, P.J. a; Antao, T.; Chang, J.T.; Chapman, B. a; Cox, C.J.; Dalke, A.; Friedberg, I.; Hamelryck, T.; Kauff, F.; Wilczynski, B.; de Hoon, M.J.L. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*. 2009; 25:1422–3. doi: 10.1093/bioinformatics/btp163. [PubMed: 19304878]
- Coenjaerts FEJ, Hoepelman AIM, Scharringa J, Aarts M, Ellerbroek PM, Bevaart L, Van Strijp J. a G. Janbon G. The Skn7 response regulator of *Cryptococcus neoformans* is involved in oxidative stress signalling and augments intracellular survival in endothelium. *FEMS Yeast Res*. 2006; 6:652–61. doi:10.1111/j.1567-1364.2006.00065.x. [PubMed: 16696662]
- De Groot MJA, Bundock P, Hooykaas PJJ, Beijersbergen AGM. *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. *Nat. Biotechnol*. 1998; 16:839–842. [PubMed: 9743116]
- Devon RS, Porteous DJ, Brookes AJ. Splinkerettes-improved vectorettes for greater efficiency in PCR walking. *Nucleic Acids Res*. 1995; 23:1644–5. [PubMed: 7784225]

- Dou X, Wang Q, Qi Z, Song W, Wang W, Guo M, Zhang H, Zhang Z, Wang P, Zheng X. MoVam7, a conserved SNARE involved in vacuole assembly, is required for growth, endocytosis, ROS accumulation, and pathogenesis of *Magnaporthe oryzae*. PLoS One. 2011; 6:e16439. doi:10.1371/journal.pone.0016439. [PubMed: 21283626]
- Feretziaki M, Heitman J. Genetic circuits that Govern Bisexual and Unisexual Reproduction in *Cryptococcus neoformans*. PLoS Genet. 2013; 9:e1003688. doi:10.1371/journal.pgen.1003688. [PubMed: 23966871]
- Fu J, Mares C, Lizcano A, Liu Y, Wickes BL. Insertional mutagenesis combined with an inducible filamentation phenotype reveals a conserved *STE50* homologue in *Cryptococcus neoformans* that is required for monokaryotic fruiting and sexual reproduction. Mol. Microbiol. 2011; 79:990–1007. doi:10.1111/j.1365-2958.2010.07501.x. [PubMed: 21299652]
- Gawronski JD, Wong SMS, Giannoukos G, Ward DV, Akerley BJ. Tracking insertion mutants within libraries by deep sequencing and a genome-wide screen for *Haemophilus* genes required in the lung. Proc. Natl. Acad. Sci. 2009; 106:16422–7. doi:10.1073/pnas.0906627106. [PubMed: 19805314]
- Gheysen G, Villarroel R, Van Montagu M. Illegitimate recombination in plants: a model for T-DNA integration. Genes Dev. 1991; 5:287–297. doi:10.1101/gad.5.2.287. [PubMed: 1995418]
- Goldman D, Vicencio A. The Chitin Connection. MBio. 2012;3. doi:10.1128/mBio.00056-12.Updated.
- Goodman AL, McNulty NP, Zhao Y, Leip D, Mitra RD, Lozupone C. a, Knight R, Gordon JI. Identifying Genetic Determinants Needed to Establish a Human Gut Symbiont in Its Habitat. Cell Host Microbe. 2009; 6:279–89. doi:10.1016/j.chom.2009.08.003. [PubMed: 19748469]
- Heger A, Belgrad TG, Goodson M. 2014 pysam documentation.
- Hu G, Caza M, Cadieux B, Chan V, Liu V, Kronstad J. *Cryptococcus neoformans* Requires the ESCRT Protein Vps23 for Iron Acquisition from Heme, for Capsule Formation, and for Virulence. Infect. Immun. 2013; 81:292–302. doi:10.1128/IAI.01037-12. [PubMed: 23132495]
- Idnurm A, Reedy J, Nussbaum J, Heitman J. *Cryptococcus neoformans* Virulence Gene Discovery through Insertional Mutagenesis. Eukaryot. Cell. 2004; 3:420–429. doi:10.1128/EC.3.2.420. [PubMed: 15075272]
- Idnurm A, Walton FJ, Floyd A, Reedy JL, Heitman J. Identification of *ENAI* as a Virulence Gene of the Human Pathogenic Fungus *Cryptococcus neoformans* through Signature-Tagged Insertional Mutagenesis. Eukaryot. Cell. 2009; 8:315–26. doi:10.1128/EC.00375-08. [PubMed: 19151325]
- Janbon G, Ormerod KL, Paulet D, Byrnes EJ, Yadav V, Chatterjee G, Mullapudi N, Hon C-C, Billmyre RB, Brunel F, Bahn Y-S, Chen W, Chen Y, Chow EWL, Coppée J-Y, Floyd-Averette A, Gaillardin C, Gerik KJ, Goldberg J, Gonzalez-Hilarion S, Gujja S, Hamlin JL, Hsueh Y-P, Ianiri G, Jones S, Kodira CD, Kozubowski L, Lam W, Marra M, Mesner LD, Mieczkowski P. a, Moyrand F, Nielsen K, Proux C, Rossignol T, Schein JE, Sun S, Wollschlaeger C, Wood I. a, Zeng Q, Neuvéglise C, Newlon CS, Perfect JR, Lodge JK, Idnurm A, Stajich JE, Kronstad JW, Sanyal K, Heitman J, Fraser J. a, Cuomo C. a, Dietrich FS. Analysis of the Genome and Transcriptome of *Cryptococcus neoformans* var. *grubii* Reveals Complex RNA Expression and Microevolution Leading to Virulence Attenuation. PLoS Genet. 2014; 10:e1004261. doi:10.1371/journal.pgen.1004261. [PubMed: 24743168]
- Jung K-W, Yang D-H, Maeng S, Lee K-T, So Y-S, Hong J, Choi J, Byun H-J, Kim H, Bang S, Song M-H, Lee J-W, Kim MS, Kim S-Y, Ji J-H, Park G, Kwon H, Cha S, Meyers GL, Wang LL, Jang J, Janbon G, Adedoyin G, Kim T, Averette AK, Heitman J, Cheong E, Lee Y-H, Lee Y-W, Bahn Y-S. Systematic functional profiling of transcription factor networks in *Cryptococcus neoformans*. Nat. Commun. 2015; 6:6757. doi:10.1038/ncomms7757. [PubMed: 25849373]
- Kemski MM, Stevens B, Rappleye C. a. Spectrum of T-DNA integrations for insertional mutagenesis of *Histoplasma capsulatum*. Fungal Biol. 2013; 117:41–51. doi:10.1016/j.funbio.2012.11.004. [PubMed: 23332832]
- Ko Y-J, Yu YM, Kim G-B, Lee G-W, Maeng PJ, Kim S, Floyd A, Heitman J, Bahn Y-S. Remodeling of global transcription patterns of *Cryptococcus neoformans* genes mediated by the stress-activated HOG signaling pathways. Eukaryot. Cell. 2009; 8:1197–217. doi:10.1128/EC.00120-09. [PubMed: 19542307]
- Kwon-Chung, KJ.; Bennett, JE. Medical Mycology. Lea & Febiger; Malvern, Pa: 1992.

- Langridge GC, Phan M-D, Turner DJ, Perkins TT, Parts L, Haase J, Charles I, Maskell DJ, Peters SE, Dougan G, Wain J, Parkhill J, Turner AK. Simultaneous assay of every *Salmonella* Typhi gene using one million transposon mutants. *Genome Res.* 2009; 19:2308–16. doi:10.1101/gr.097097.109. [PubMed: 19826075]
- Leinonen R, Sugawara H, Shumway M. The Sequence Read Archive. *Nucleic Acids Res.* 2011; 39:D19–21. doi:10.1093/nar/gkq1019. [PubMed: 21062823]
- Leoni C, Volpicella M, De Leo F, Gallerani R, Ceci LR. Genome walking in eukaryotes. *FEBS J.* 2011; 278:3953–77. doi:10.1111/j.1742-4658.2011.08307.x. [PubMed: 21848672]
- Lesage G, Shapiro J, Specht C. a, Sdicu A-M, Ménard P, Hussein S, Tong AHY, Boone C, Bussey H. An interactional network of genes involved in chitin synthesis in *Saccharomyces cerevisiae*. *BMC Genet.* 2005;6. doi:10.1186/1471-2156-6-8. [PubMed: 15701179]
- Li G, Zhou Z, Liu G, Zheng F, He C. Characterization of T-DNA insertion patterns in the genome of rice blast fungus *Magnaporthe oryzae*. *Curr. Genet.* 2007; 51:233–43. doi:10.1007/s00294-007-0122-5. [PubMed: 17372735]
- Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv Prepr.* 2013; 00:1–3.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009; 25:2078–9. doi:10.1093/bioinformatics/btp352. [PubMed: 19505943]
- Lin X, Jackson JC, Feretzaki M, Xue C, Heitman J. Transcription Factors Mat2 and Znf2 Operate Cellular Circuits Orchestrating Opposite- and Same-Sex Mating in *Cryptococcus neoformans*. *PLoS Genet.* 2010; 6:e1000953. doi:10.1371/journal.pgen.1000953. [PubMed: 20485569]
- Maeng S, Ko Y-J, Kim G-B, Jung K-W, Floyd A, Heitman J, Bahn Y-S. Comparative transcriptome analysis reveals novel roles of the Ras and cyclic AMP signaling pathways in environmental stress response and antifungal drug sensitivity in *Cryptococcus neoformans*. *Eukaryot. Cell.* 2010; 9:360–78. doi:10.1128/EC.00309-09. [PubMed: 20097740]
- Mayerhofer R, Koncz-Kalman Z, Nawrath C, Bakkeren G, Cramer A, Angelis K, Redei GP, Schell J, Hohn B, Koncz C. T-DNA integration: a mode of illegitimate recombination in plants. *EMBO J.* 1991; 10:697–704. [PubMed: 2001683]
- McDade HC, Cox GM. A new dominant selectable marker for use in *Cryptococcus neoformans*. *Med. Mycol.* 2001; 39:151–4. [PubMed: 11270405]
- Meng Y, Patel G, Heist M, Betts MF, Tucker SL, Galadima N, Donofrio NM, Brown D, Mitchell TK, Li L, Xu J-R, Orbach M, Thon M, Dean RA, Farman ML. A systematic analysis of T-DNA insertion events in *Magnaporthe oryzae*. *Fungal Genet. Biol.* 2007; 44:1050–64. doi:10.1016/j.fgb.2007.04.002. [PubMed: 17544743]
- Michiels CB, Hooykaas PJJ, van den Hondel C. a M.J.J. Ram AFJ. *Agrobacterium*-mediated transformation as a tool for functional genomics in fungi. *Curr. Genet.* 2005; 48:1–17. doi: 10.1007/s00294-005-0578-0. [PubMed: 15889258]
- Mullins E, Chen X, Romaine P, Raina R, Geiser D, Kang S. *Agrobacterium*-Mediated Transformation of *Fusarium oxysporum*: An Efficient Tool for Insertional Mutagenesis and Gene Transfer. *Phytopathology.* 2001; 91:173–180. [PubMed: 18944391]
- O'Meara TR, Holmer SM, Selvig K, Dietrich F, Alspaugh JA. *Cryptococcus neoformans* Rim101 Is Associated with Cell Wall Remodeling and Evasion of the Host Immune Responses. *MBio.* 2013; 4:1–13. doi:10.1128/mBio.00522-12.Editor.
- O'Meara TR, Norton D, Price MS, Hay C, Clements MF, Nichols CB, Alspaugh JA. Interaction of *Cryptococcus neoformans* Rim101 and Protein Kinase A Regulates Capsule. *PloS Pathog.* 2010;6. doi:10.1371/journal.ppat.1000776.
- O'Meara TR, Xu W, Selvig KM, O'Meara MJ, Mitchell AP, Alspaugh JA. The *Cryptococcus neoformans* Rim101 Transcription Factor Directly Regulates Genes Required for Adaptation to the Host. *Mol. Cell. Biol.* 2014; 34:673–84. doi:10.1128/MCB.01359-13. [PubMed: 24324006]
- Ohashi K, Kawai S, Murata K. Secretion of quinolinic acid, an intermediate in the kynurenine pathway, for utilization in NAD⁺ biosynthesis in the yeast *Saccharomyces cerevisiae*. *Eukaryot. Cell.* 2013; 12:648–653. doi:10.1128/EC.00339-12. [PubMed: 23457190]

- Ost KS, O'Meara TR, Huda N, Esher SK, Alspaugh JA. The *Cryptococcus neoformans* Alkaline Response Pathway: Identification of a Novel Rim Pathway Activator. *PLoS Genet.* 2015; 11:e1005159. doi:10.1371/journal.pgen.1005159. [PubMed: 25859664]
- Panozzo C, Nawara M, Suski C, Kucharczyka R, Skoneczny M, Bécam AM, Rytka J, Herbert CJ. Aerobic and anaerobic NAD⁺ metabolism in *Saccharomyces cerevisiae*. *FEBS Lett.* 2002; 517:97–102. doi:10.1016/S0014-5793(02)02585-1. [PubMed: 12062417]
- Park BJ, Wannemuehler K. a, Marston BJ, Govender N, Pappas PG, Chiller TM. Estimation of the current global burden of cryptococcal meningitis among persons living with HIV/AIDS. *AIDS.* 2009; 23:525–30. doi:10.1097/QAD.0b013e328322ffac. [PubMed: 19182676]
- Peng RD. Reproducible research and *Biostatistics*. *Biostatistics.* 2009; 10:405–8. doi:10.1093/biostatistics/kxp014. [PubMed: 19535325]
- Perfect JR, Lang SDR, Durack DT. Chronic Cryptococcal Meningitis: A New Experimental Model in Rabbits. *Am. J. Pathol.* 1980; 101:177–194. [PubMed: 7004196]
- Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. *Nat. Biotechnol.* 2011; 29:24–26. doi:10.1038/nbt0111-24. [PubMed: 21221095]
- Rodrigues ML, Nakayasu ES, Oliveira DL, Nimrichter L, Nosanchuk JD, Almeida IC, Casadevall A. Extracellular vesicles produced by *Cryptococcus neoformans* contain protein components associated with virulence. *Eukaryot. Cell.* 2008a; 7:58–67. doi:10.1128/EC.00370-07. [PubMed: 18039940]
- Rodrigues ML, Nimrichter L, Oliveira DL, Nosanchuk JD, Casadevall A. Vesicular Trans-Cell Wall Transport in Fungi: A Mechanism for the Delivery of Virulence-Associated Macromolecules? *Lipid Insights.* 2008b; 2:27–40. [PubMed: 20617119]
- Sato T, Darsow T, Emr S. Vam7p , a SNAP-25-Like Molecule , and Vam3p , a Syntaxin Homolog , Function Together in Yeast Vacuolar Protein Trafficking. *Mol. Cell. Biol.* 1998:18.
- Sharma CB, Knauer R, Lehle L. Biosynthesis of lipid-linked oligosaccharides in yeast: The *ALG3* gene encodes the Dol-P-Man:Man₅GlcNAc₂-PP-Dol Mannosyltransferase. *Biol. Chem.* 2001; 382:321–328. doi:10.1515/BC.2001.039. [PubMed: 11308030]
- Sherman F. Getting Started with Yeast. *Methods Enzym.* 1991; 194:3–21.
- Shimizu K, Imanishi Y, Toh-E A, Uno J, Chibana H, Hull CM, Kawamoto S. Functional characterization of *PMT2*, encoding a protein-*O*-mannosyltransferase, in the human pathogen *Cryptococcus neoformans*. *Fungal Genet. Biol.* 2014; 69:13–22. doi:10.1016/j.fgb.2014.05.007. [PubMed: 24892554]
- Sugui JA, Chang YC, Kwon-Chung KJ. *Agrobacterium tumefaciens*-Mediated Transformation of *Aspergillus fumigatus*: an Efficient Tool for Insertional Mutagenesis and Targeted Gene Disruption. *Appl. Environ. Microbiol.* 2005; 71:1798–1802. doi:10.1128/AEM.71.4.1798. [PubMed: 15812003]
- Sullivan T, Rooney P, Klein B. *Agrobacterium tumefaciens* Integrates Transfer DNA into Single Chromosomal Sites of Dimorphic Fungi and Yields Homokaryotic Progeny from Multinucleate Yeast. *Eukaryot. Cell.* 2002; 1:895–905. doi:10.1128/EC.1.6.895. [PubMed: 12477790]
- Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* 2013; 14:178–92. doi:10.1093/bib/bbs017. [PubMed: 22517427]
- Toffaletti D, Rude T, Johnston S, Durack D, Perfect J. Gene transfer in *Cryptococcus neoformans* by use of biolistic delivery of DNA. *J. Bacteriol.* 1993; 175:1405–1411. [PubMed: 8444802]
- Triglia T, Peterson MG, Kemp DJ. A procedure for in vitro amplification of DNA segments that lie outside the boundaries of known sequences. *Nucleic Acids Res.* 1988; 16:8186. [PubMed: 3047679]
- Ungermann C, Wickner W. Vam7p, a vacuolar SNAP-25 homolog, is required for SNARE complex integrity and vacuole docking and fusion. *EMBO J.* 1998; 17:3269–76. doi:10.1093/emboj/17.12.3269. [PubMed: 9628864]
- Van der Walt S, Colbert SC, Varoquaux G. The NumPy array: a structure for efficient numerical computation. *Comput. Sci. Eng.* 2011; 13:22–30.

- Van Opijnen T, Bodi KL, Camilli A. Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nat. Methods*. 2009; 6:767–72. doi:10.1038/nmeth.1377. [PubMed: 19767758]
- Vega K, Kalkum M. Chitin, Chitinase Responses, and Invasive Fungal Infections. *Int. J. Microbiol.* 2012;2012:920459. doi:10.1155/2012/920459. [PubMed: 22187561]
- Wada Y, Anraku Y. Genes for Directing Vacuolar Morphogenesis in *Saccharomyces cerevisiae* II. *VAM7*, A Gene for Regulating Morphogenic Assembly of the Vacuoles. *J. Biol. Chem.* 1992; 267:18671–18675. [PubMed: 1526999]
- Wada Y, Ohsumi Y, Anraku Y. Genes for Directing Vacuolar Morphogenesis in *Saccharomyces cerevisiae* I. Isolation and Characterization of Two Classes of Vam Mutants. *J. Biol. Chem.* 1992; 267:18665–18670. [PubMed: 1526998]
- Wagener J, Malireddi RKS, Lenardon MD, Köberle M, Vautier S, MacCallum DM, Biedermann T, Schaller M, Netea MG, Kanneganti T-D, Brown GD, Brown AJP, Gow NAR. Fungal Chitin Dampens Inflammation through IL-10 Induction Mediated by NOD2 and TLR9 Activation. *PLoS Pathog.* 2014; 10:e1004050. doi:10.1371/journal.ppat.1004050. [PubMed: 24722226]
- Walton FJ, Idnurm A, Heitman J. Novel gene functions required for melanization of the human pathogen *Cryptococcus neoformans*. *Mol. Microbiol.* 2005; 57:1381–96. doi:10.1111/j.1365-2958.2005.04779.x. [PubMed: 16102007]
- Wedlich-Söldner R, Bölker M, Kahmann R, Steinberg G. A putative endosomal t-SNARE links exo- and endocytosis in the phytopathogenic fungus *Ustilago maydis*. *EMBO J.* 2000; 19:1974–1986. doi:10.1093/emboj/19.9.1974. [PubMed: 10790364]
- Wolf JM, Espadas-Moreno J, Luque-Garcia JL, Casadevall A. Interaction of *Cryptococcus neoformans* extracellular vesicles with the cell wall. *Eukaryot. Cell.* 2014; 13:1484–93. doi:10.1128/EC.00111-14. [PubMed: 24906412]
- Wormley F, Heinrich G, Miller J, Perfect J, Cox G. Identification and characterization of an *SKN7* homologue in *Cryptococcus neoformans*. *Infect. Immun.* 2005; 73:1–10. doi:10.1128/IAI.73.8.5022.
- Youseff BH, Dougherty JA, Rappleye CA. Reverse genetics through random mutagenesis in *Histoplasma capsulatum*. *BMC Microbiol.* 2009; 9:236. doi:10.1186/1471-2180-9-236. [PubMed: 19919692]
- Zhang H, Li B, Fang Q, Li Y, Zheng X, Zhang Z. SNARE protein FgVam7 controls growth, asexual and sexual development, and plant infection in *Fusarium graminearum*. *Mol. Plant Pathol.* 2015 doi:10.1111/mpp.12267.

Highlights

- We screened a library of *C. neoformans* insertional mutants for cell wall phenotypes
- We developed a sequencing/analysis pipeline to identify sites of random insertion
- We identified & confirmed insertion sites from a pool of *C. neoformans* cell wall mutants
- Our analysis also revealed atypical insertion events associated with AMT
- AIM-Seq can be immediately applied to a range of organisms and mutagenesis methods

Figure 1. The AIM-Seq Method

A library of insertional mutants is generated by *Agrobacterium*-mediated transconjugation, followed by a genetic screen for phenotypes of interest. Genomic DNA from mutants of interest is pooled and sequenced. Sequence data is analyzed by AIM-HII software, which (1) extracts those containing T-DNA-genomic junctions, (2) trims the T-DNA sequence, and (3) maps the remaining sequence to the annotated genome. AIM-HII identifies clusters of junction reads, and then matches clusters corresponding to the ends of the insertion.

Figure 2. Insertion sites identified by AIM-HII

(a). Twenty-seven mutants of interest were pooled and submitted for whole genome sequencing and analysis. AIM-HII identified 35 putative insertion sites. Sixteen sites displayed typical insertion characteristics with clustered reads on both sides of a putative insert. Fourteen of the 16 displayed no deletion, or short insertion-induced deletions, while 2 inserts induced larger deletions. Nineteen sites were singletons, with clustered reads on only one side of a putative insertion, suggesting chromosomal rearrangements. (b). A typical insertion site with reads clustering on both sides of a putative insert-induced deletion. Region where sequences flanking an insert align is shown in the expanded view. Grey boxes represent exons. (c). An insertion-induced large deletion of ~1.2 kb. The deleted region, along with regions where sequences flanking the insert align to the genome, are shown in the expanded view (d). A singleton insertion with reads clustering on one side of a putative T-DNA insert. Expanded view shows region where sequences flanking one side of an insert align to the genome.

Figure 3. T-DNA insert characteristics

T-DNA sequences extracted from whole genome sequence data aligned to the full-length pPZP-NATcc Ti plasmid. The depth (number) of reads mapping to the full-length plasmid demonstrates that most of the inserted plasmid DNA is from the T-DNA insert portion.

However the uniform appearance of reads across the backbone clearly indicates at least one instance in which the plasmid backbone has been incorporated.

Figure 4. PCR and phenotypic validation of an insertion in the *C. neoformans* *ACA1* gene (CNAG_05218) identified by AIM-Seq

(a). Plot of the insertion site in the open reading frame of *ACA1*. Expanded view shows the region where insert flanking sequences align. Grey boxes indicate exons. (b). PCR confirmation of this site in strain WHM24. Primers flanking the putative insertion site were used to differentiate between wild type and mutated loci in AMT mutants. In lane 3, WHM24 shows a shift to a larger amplification product of 2.8 kb relative to the WT 500 bp product in lane 2. HyperLadder, 1kb, shown in lane 1 for reference (Bioline). (c). Phenotypic validation of WHM24. The sensitivity to pH 8 and growth defect on 1.5 M NaCl of the original AMT mutant, WHM24, are identical in an independent strain with an *aca1* mutation.

Figure 5. Virulence- and cell wall-associated phenotypes of *C. neoformans* insertional mutants identified by AIM-Seq

(a). Indicated strains were normalized, serially diluted, and incubated on YPD medium with the addition of indicated cell wall perturbing agents or at the indicated temperatures. (b). Strains were incubated in capsule-inducing conditions (CO₂- independent medium, 37°C) with shaking for 72 hours. Capsule was assessed by India ink counterstaining, followed by imaging (630x). (c). Melanin production was assessed by the appearance of brown pigment in colonies incubated on Niger seed agar for 2-3 days at the indicated temperatures.

Figure 6. Cell wall chitin and chitosan exposure as assessed by staining of *C. neoformans* insertional mutants identified by AIM-Seq

Strains were incubated in CO₂- independent medium at 37°C for 18 hours prior to staining and imaging. Exposure of cell wall components was assessed by staining with (a) CFW and WGA for chitin, and (b) eosin Y for chitosan. Stained cells were imaged by fluorescent microscopy (630x).

Figure 7. Growth defect of WHM11/*cnag_02793* *T-DNA* on YNB medium can be rescued by exogenous quinolinic acid

Cells were normalized, serially diluted, and incubated on YPD or YNB medium with the addition of the indicated concentrations of quinolinic acid (QA).

Figure 8. Vacuolar staining of WHM27/*cnag_02676* *T-DNA*

Strains were incubated in YPD liquid media at 30°C for 18 hours prior to staining and imaging. Cells were stained with the endocytosis marker, FM4-64, for 10 minutes on ice, followed by washing and incubation at room temperature for 30 minutes. Stained cells were imaged by fluorescent microscopy (630x). Black arrows indicate FM4-64 staining of vacuolar membranes.

Table 1Strains used in this study^a

Strain	Genotype	Reference
H99	<i>MATa</i>	(Perfect et al., 1980)
YSB6	<i>MATa aca1 ::NAT-STM#43</i>	(Bahn et al., 2004)
YSB349	<i>MATa skn7 ::NAT-STM#201</i>	(Bahn et al., 2006)
SKE47	<i>MATa cnag_05142^{T-DNA}</i>	This study
SKE40	<i>MATa cnag_02793^{T-DNA}</i>	This study
SKE49	<i>MATa cnag_02676^{T-DNA}</i>	This study

^a AMT strains are not listed. All were generated in the H99 *MATa* background

Table 2

Primers used in this study.

Primer Name	Primer Sequence	Purpose
AA428	GGGCATGCTCATGTAGAGC	<i>NAT</i> Southern probe
AA621	CCACTCTTGACGACACGGCT	<i>NAT</i> Southern probe
AA3847	GCCACTTCCATTTTCCTTTG	WHM4 insert 5' flank
AA4008	ACGGGCAAAGAGCATACCTA	WHM4 insert 3' flank
AA3837	GTGCCCTCCAGCTGTAGTCC	WHM7 insert 5' flank
AA3838	TATCGGATGCAGATGTGTTC	WHM7 insert 3' flank
AA4009	CTGGTCATTGGATGTGATGG	WHM11 insert 5' flank, WHM11-SKE Southern probe 5'
AA4010	TAGTTTTGCAAGGCCAGTTC	WHM11 insert 3' flank
AA3833	CGGTCGAGTGTTGATCTGTG	WHM12 insert 5' flank
AA3834	TAGTGAGGGGCGATGGAGTC	WHM12 insert 3' flank
AA3928	AAATCCCGGTCAACATCAAA	WHM13 insert 5' flank
AA3929	GGATGGGGATGCTGAAGTAA	WHM13 insert 3' flank
AA3843	GAGGCCGTATAGAAGGTT	WHM15 insert 5' flank
AA3844	CATGGAGTAGGAGATGATCG	WHM15 insert 3' flank
AA4016	TGTCCGAGGAGGTGAGGTAG	WHM19 insert 5' flank
AA4017	GTTGAGAACGGAGGACCTTG	WHM19 insert 3' flank
AA3850	GCTGCTGAGAGCCTGAAGGC	WHM21 insert 5' flank
AA3851	ACCAGACCATCGAGCTGGCC	WHM21 insert 3' flank
AA3963	CTTGGGGCTGGGTACAAGT	WHM24 insert 5' flank
AA3964	TCGCAATCCAGAAGTGACAG	WHM24 insert 3' flank
AA3854	CCATACAATCCCCGAAAGAG	WHM27 insert 5' flank
AA3855	GGAGAGGAAGTCGAAAAGG	WHM27 insert 3' flank
AA4087	TTCTGTGATCCGTCCTTTC	WHM4-SKE Southern probe 5'
AA4088	AGCTTATGTCTGAGGGGTTGA	WHM4-SKE Southern probe 3'
AA4102	TACAATGTCCACTGCGGAAG	WHM11-SKE Southern probe 3'
AA4086	ACAGGGTTTGTCTATGGCTCT	WHM27-SKE Southern probe 5'
AA3927	TTCCCGTCTCCTGATTTAC	WHM27-SKE Southern probe 3'

Table 3

AIM-HII Identified Insert Locations

Chromosome	Position	Number of Reads	Insert Location	Insert Type	Gene ID
Chromosome 1	1515285-1515409	5	Intergenic	Singleton	--
Chromosome 1	1752283-1752786	45	Promoter	Pair	CNAG_00674
Chromosome 2	774969-775175	15	CDS	Singleton	CNAG_03772
Chromosome 2	991283-991766	10	CDS	Pair	CNAG_03844
Chromosome 3	596973-597253	12	CDS	Pair	CNAG_02857
Chromosome 3	754720-754915	10	CDS	Pair	CNAG_02793
Chromosome 3	1050456-1050652	19	CDS	Pair	CNAG_02676
Chromosome 3	1056712-1057010	7	Promoter	Singleton	CNAG_02675
Chromosome 3	1257168-1257416	7	Promoter	Singleton	CNAG_02601, CNAG_02602
Chromosome 4	167147-167344	15	CDS	Singleton	CNAG_04991
Chromosome 4	373261-373562	10	CDS	Singleton	CNAG_05076
Chromosome 4	541443-543070	19	CDS	Pair	CNAG_05142
Chromosome 4	759966-760537	18	CDS	Pair	CNAG_05218
Chromosome 5	219282-219737	11	CDS	Pair	CNAG_06804
Chromosome 5	556667-557021	6	CDS	Singleton	CNAG_01339, CNAG_01336
Chromosome 5	663564-663833	5	CDS	Singleton	CNAG_01294
Chromosome 5	734791 - 735049	9	CDS	Singleton	CNAG_01272
Chromosome 5	796816-797015	15	Intergenic	Singleton	CNAG_01251, CNAG_01252
Chromosome 6	804933-805064	5	Intergenic	Singleton	--
Chromosome 6	880446-880782	11	Promoter	Singleton	CNAG_02220
Chromosome 6	1165644-1165974	10	CDS	Pair	CNAG_02109
Chromosome 7	565664-565912	10	Intergenic	Singleton	--
Chromosome 8	72699-72836	12	CDS, Promoter	Singleton	CNAG_03109, CNAG_03110
Chromosome 8	885191-885606	16	CDS, Promoter	Pair	CNAG_03409, CNAG_03410
Chromosome 9	190056-190580	19	CDS	Pair	CNAG_04171
Chromosome 9	416783-416940	5	CDS, Promoter	Singleton	CNAG_04256, CNAG_04257
Chromosome 9	1133822-1134115	10	CDS	Singleton	CNAG_07780
Chromosome 10	770026-770300	61	Intergenic	Pair	CNAG_04653, CNAG_04654
Chromosome 11	262855-263019	15	CDS	Singleton	CNAG_01552
Chromosome 12	257752-258181	13	Intergenic	Pair	--
Chromosome 13	266250-266756	8	CDS	Pair	CNAG_06348
Chromosome 13	429584-430045	13	CDS	Pair	CNAG_06418, CNAG_06419
Chromosome 14	151923-152160	8	Promoter	Singleton	CNAG_05381
Chromosome 14	340727-341091	8	Promoter, CDS	Pair	CNAG_05449, CNAG_07876
Chromosome 14	428615-428878	18	Intergenic	Singleton	--

Table 4

AMT strains positively matched to insertion sites by PCR

AMT Strain ID	Insert Location	Gene ID
WHM4	CDS	CNAG_05142
WHM11	CDS	CNAG_02793
WHM12	Promoter	CNAG_00674
WHM13	CDS	CNAG_02857
WHM15	Intergenic	CNAG_04653, CNAG_04654
WHM19	CDS	CNAG_03844
WHM21	CDS, Promoter	CNAG_03409, CNAG_03410
WHM24	CDS	CNAG_05218
WHM27	CDS	CNAG_02676

Table 5

AMT strains phenotypically validated in independent mutant strains

AMT Strain ID	Independent Disruption	Gene ID	Annotation	Phenotypes ^b		
				pH 8	1.5 M NaCl	0.06% SDS
WHM4	--	CNAG_05142	alpha-1,3-mannosyltransferase	+	±	–
--	WHM4-SKE ^a			+	±	–
WHM11	--	CNAG_02793	kynurenine 3-monooxygenase	±	+	–
--	WHM11-SKE			±	+	–
WHM21	--	CNAG_03409, CNAG_03410	Skn7, cytoplasmic protein	+	±	±
--	<i>skn7</i>			+	±	+
WHM24	--	CNAG_05218	adenylyl cyclase-associated protein	±	–	–
--	<i>aca1</i>			±	–	–
WHM27	--	CNAG_02676	conserved hypothetical protein	+	–	–
--	WHM27-SKE			+	–	–

^aGenerated by disruption with original T-DNA insert in the H99 *MATa* background^b+: robust growth, ±: moderate growth/sensitivity, –: no growth